PEAN: 3D Hand Pose Estimation Adversarial Network

Linhui Sun[1,2], Yifan Zhang[1,2], Jian Cheng[1,2], Hanqing Lu[1,2]
[1]NLPR & AIRIA, Institute of Automation, Chinese Academy of Sciences
[2]School of Artificial Intelligence, University of Chinese Academy of Sciences

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Method

3D hand pose estimation still suffers from the problems of predicting inaccurate or invalid poses which conflict with physical and kinematic constraints. To address these problems, we propose a novel 3D hand pose estimation adversarial network (PEAN) which can implicitly utilize such constraints to regularize the prediction in an adversarial learning framework.

PEAN contains two parts: a 3D hierarchical estimation network (3DHNet) to predict hand pose, which decouples the task into multiple subtasks with a hierarchical structure; a pose discrimination network (PDNet) to judge the reasonableness of the estimated 3D hand pose, which back-propagates the constraints to the estimation network.

During the adversarial learning process, PDNet is expected to distinguish the estimated 3D hand pose and the ground truth, while 3DHNet is expected to estimate more valid pose to confuse PDNet.

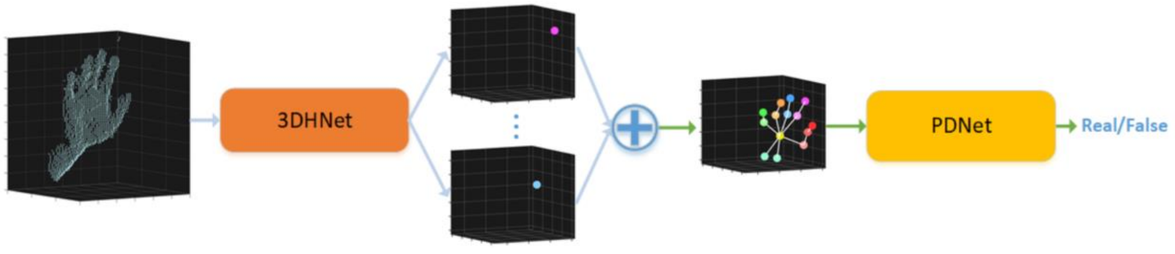

Fig. 1. Overall structure of PEAN. PEAN is comprised of two components: a hierarchical hand pose estimation network, named 3DHNet, and a pose discrimination network, named PDNet.

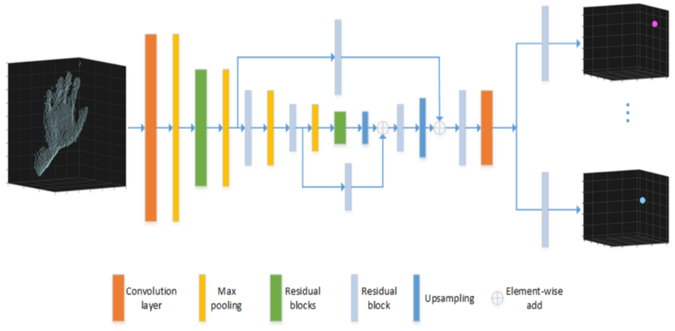

Fig. 2. Network structure of 3DHNet. Based on 3D CNN, the input of 3DHNet is the volumetric form of hand and the output is 3D heatmap of each joint.

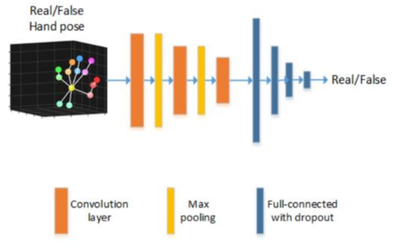

Fig. 3. Network structure of PDNet. The input of PDNet is 3D hand pose formed by ground truth or formed by the predicted results of 3DHNet, which will be judged as real or false pose by PDNet.

Results

The proposed PEAN achieves the state-of-the-art results on three challenging hand pose estimation datasets: NYU, MSRA and ICVL.

Methods	Average 3D distance error / mm		
	ICVL	NYU	MSRA
Feedback	-	-	15.97
Lie-X	-	-	14.51
DeepPrior	10.4	19.73	-
DeepPrior++	8.1	12.24	9.5
DeepModel	11.56	17.04	-
Pose-Ren	6.79	11.81	8.65
Ren-4x6x6	7.63	13.39	-
Ren-9x6x6	7.31	12.69	9.7
DensReg	7.3	10.27	**7.2**
CrossInfoNet	6.73	10.08	7.86
3DCNN	-	14.1	9.6
SHPR-Net	7.22	10.78	7.86
HandPointNet	6.94	10.54	8.5
Point-to-Point	6.3	9.1	7.7
V2V-PoseNet*	**6.28**	**8.42**	7.59
3DHNet(ours)	**6.28**	**8.73**	**7.58**
PEAN(ours)	**5.85**	**8.59**	**6.91**

The results of ablation study verify the ability of 3DHNet .

Structures	Average 3D distance error / mm		
	NYU	MSRA	ICVL
Base	8.73	7.58	6.28
Base+PDNet	8.59	6.91	5.85